New Advances: Global Assimilation and Modeling

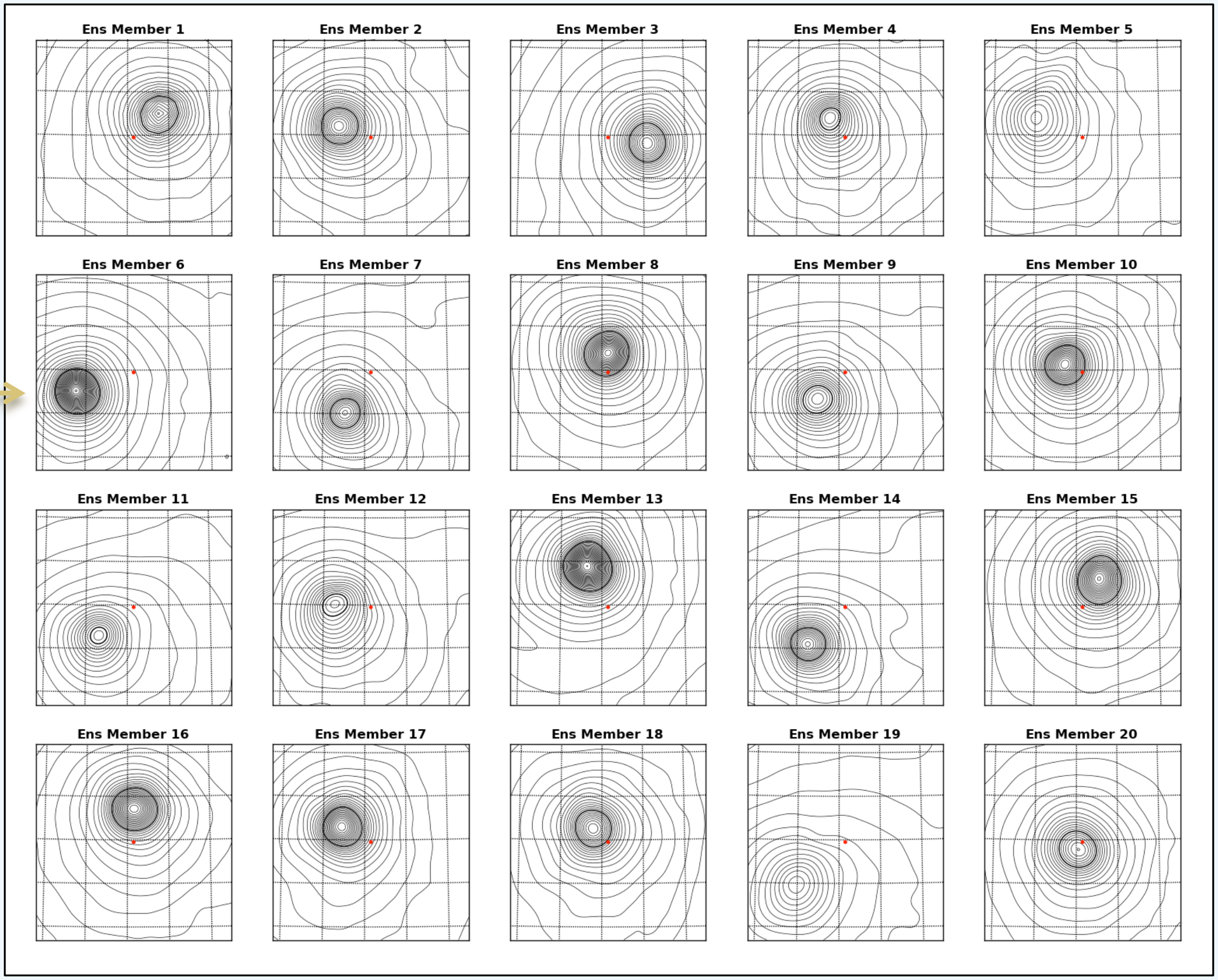
Jeffrey Whitaker

Earth System Research Laboratory

The Ensemble Kalman Filter (EnKF): Weather and Climate Applications

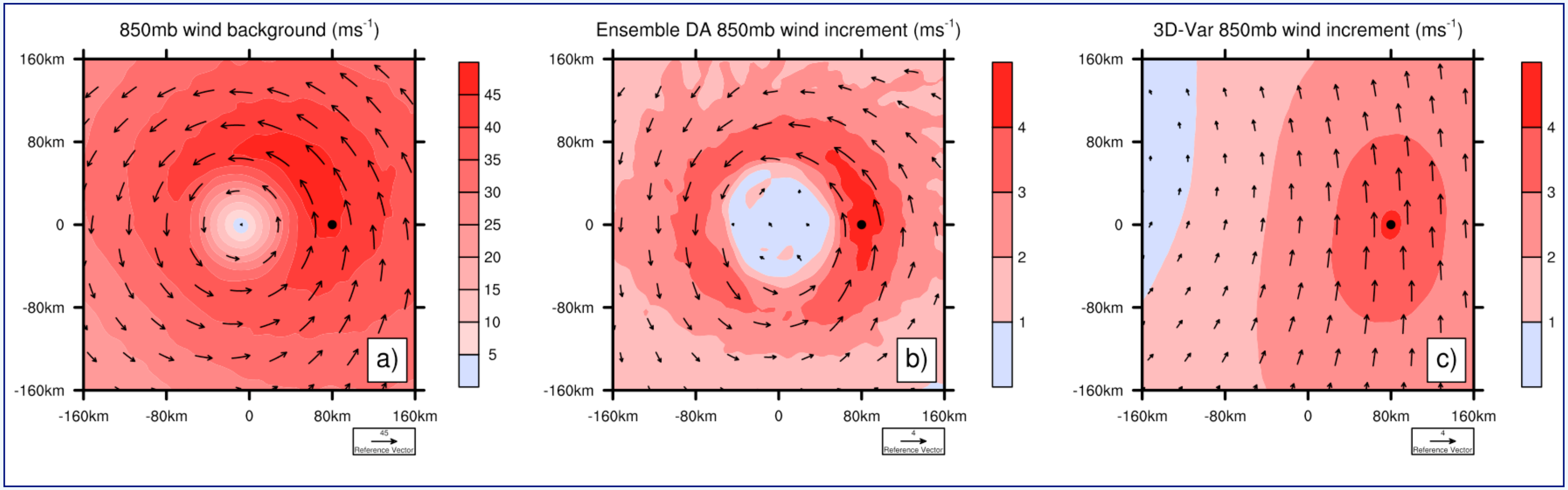
Ensembles Provide Estimates of Forecast Error & Their Correlation Structure

20-Member Ensemble Of Short-Term Forecasts, Showing Uncertainty In the Forecast Position and Structure of a Hurricane Vortex



The Ensemble Kalman Filter (EnKF)

- A method for improving data assimilation that uses uncertainty estimates from an ensemble (and produces an ensemble of analyses)

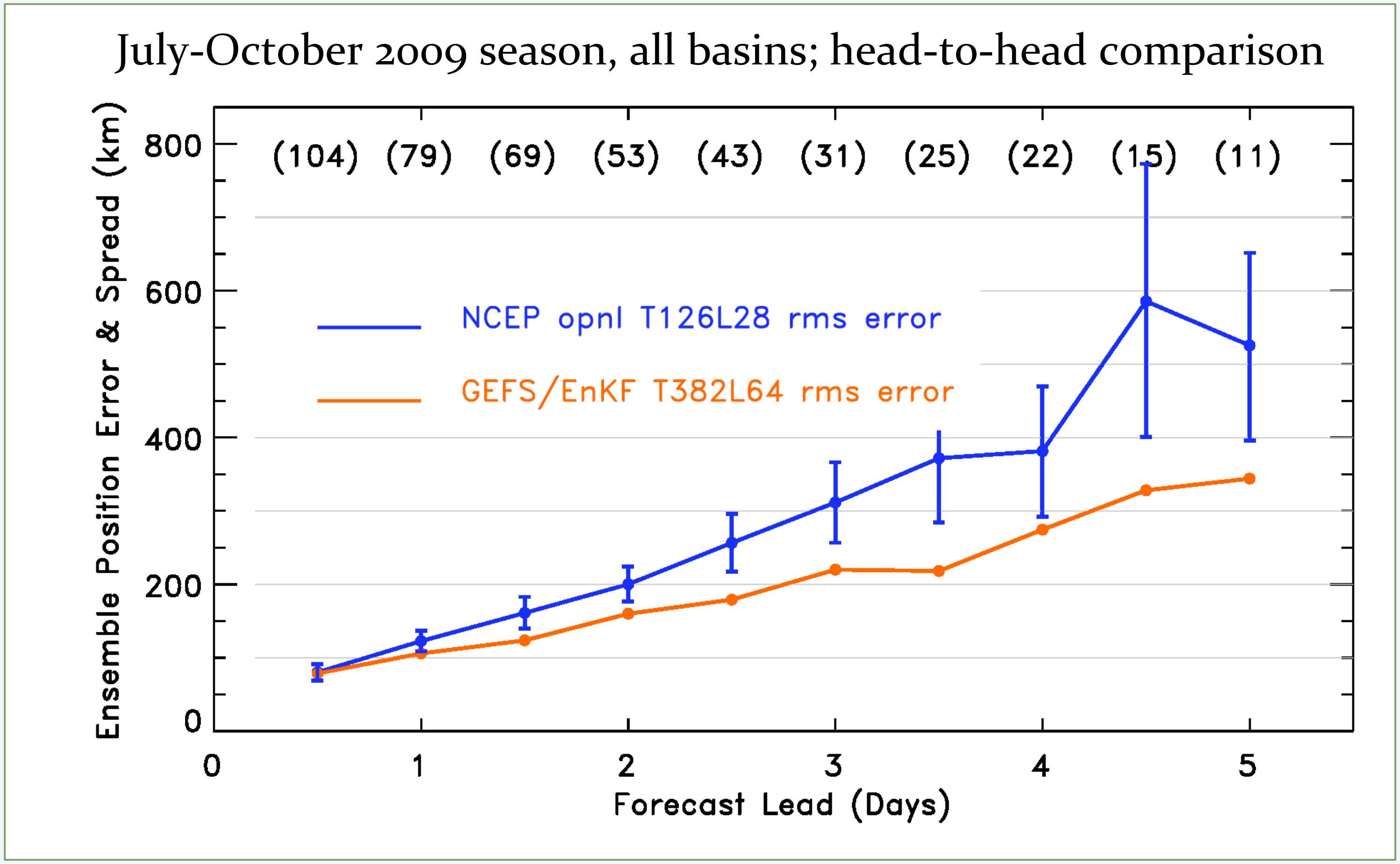


This shows the adjustment to a wind observation 1 m/s greater to the background (at dot) in EnKF and in more classical "3D-Var"

figure c/o Xuguang Wang, formerly CIRES/ESRL, now University of Oklahoma

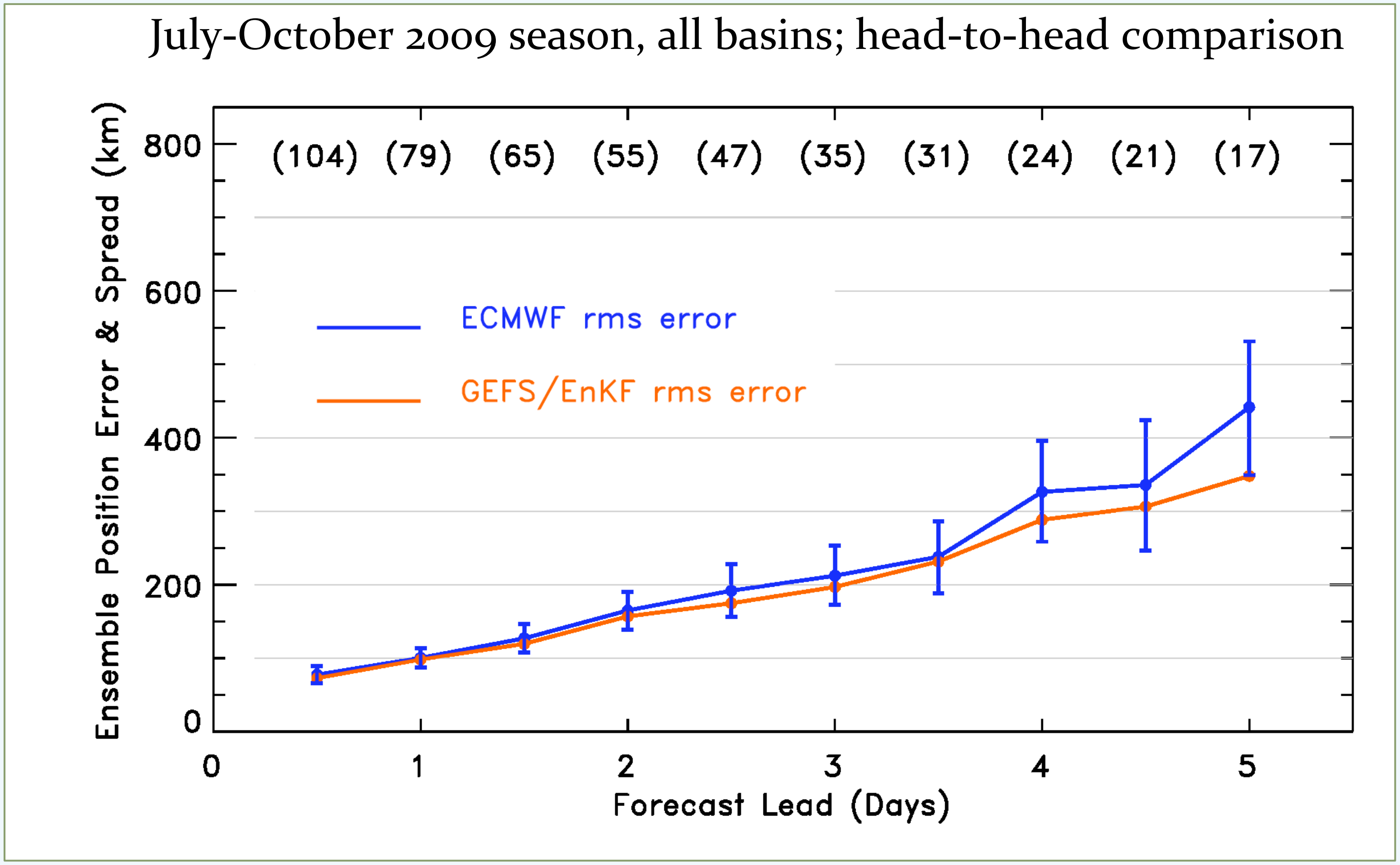
Part 1:
Improving Weather Forecasts

Track Errors, High-Resolution GFS/EnKF vs. Lower-Resolution Operational GFS/GSI-ET



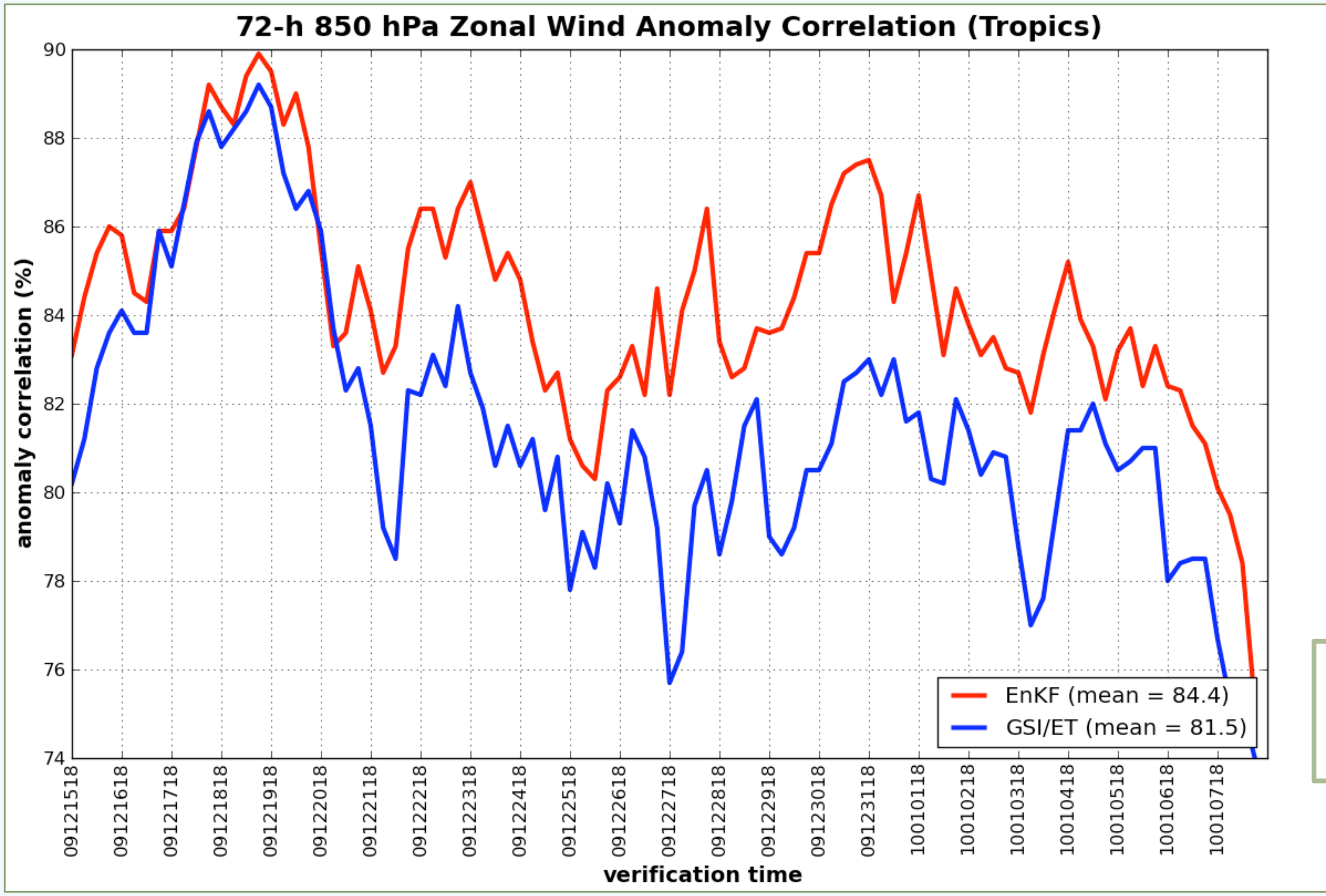
higher resolution + EnKF = clear improvement

Track Errors, High-Resolution GFS/EnKF vs. Operational ECMWF Ensemble



Competitive with ECMWF for Track Error (Comparable Resolution)

Anomaly Correlation for Mean of EnKF Ensemble vs. NCEP Operational GSI + ET Perturbations



Here, 20-member ensemble, T190 GFS used by both

Large improvement in tropical wind scores – better TC steering?

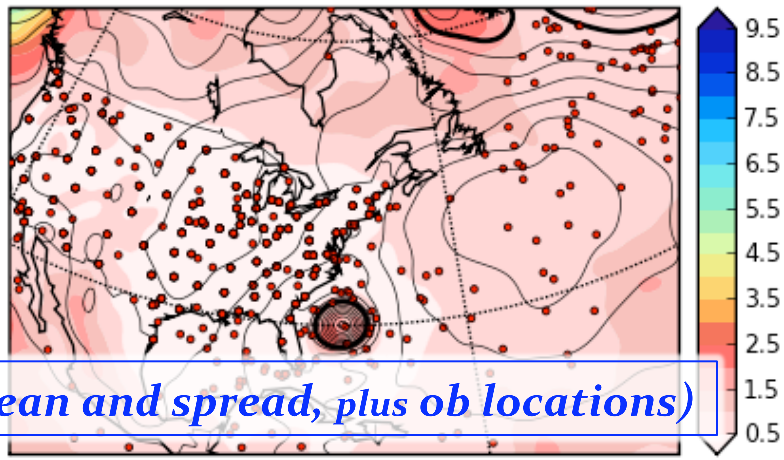
EnKF Partnership with NCEP/EMC & NASA/GMAO

- Test for possible future operational implementation at NCEP/EMC:
  - Future hybrid variational / EnKF system.
  - Replacement for current “ensemble transform” method of generating initial ensemble perturbations around control forecast.

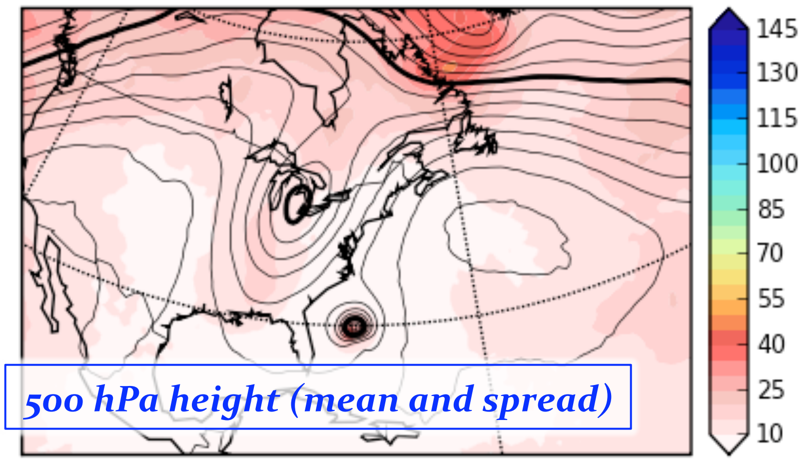
Part 2:
Climate Applications

Climate Reanalysis with Sparse Data

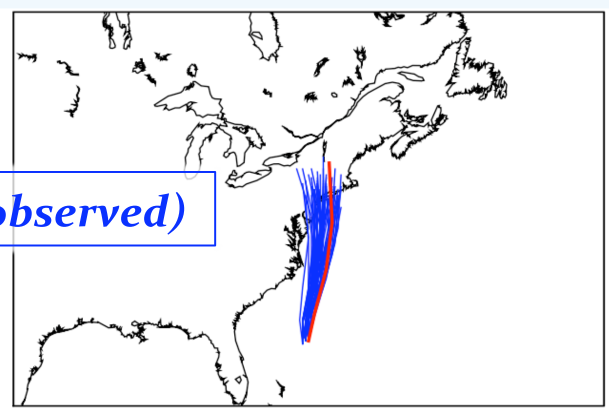
Reanalysis of the 1938 New England Hurricane using only p_s obs



MSLP (mean and spread, plus ob locations)



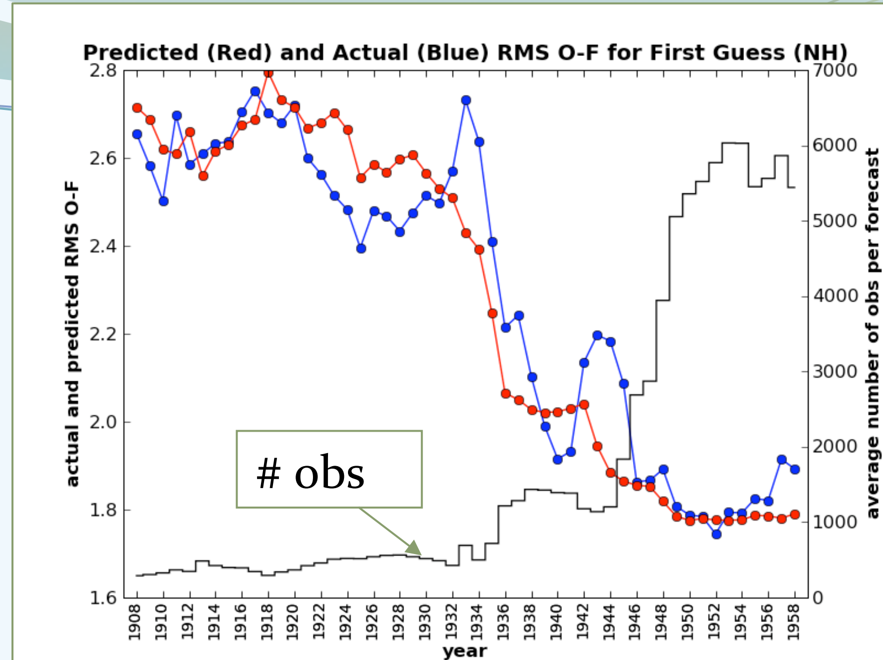
500 hPa height (mean and spread)

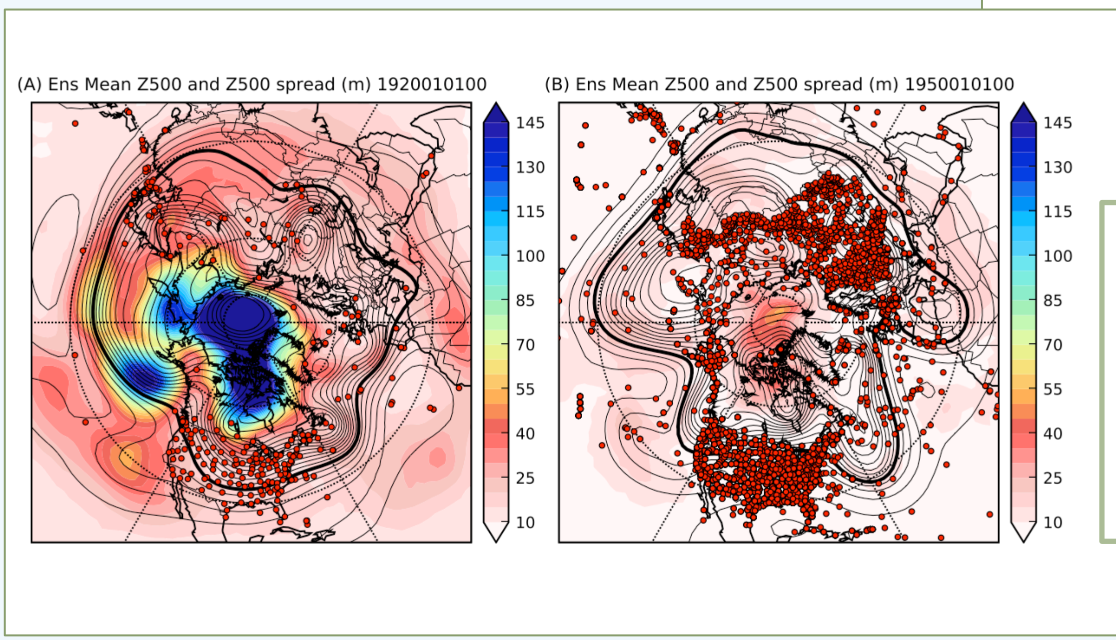


24-h track forecasts (red observed)

Estimating Space and Time-Varying Uncertainty in Reanalyses

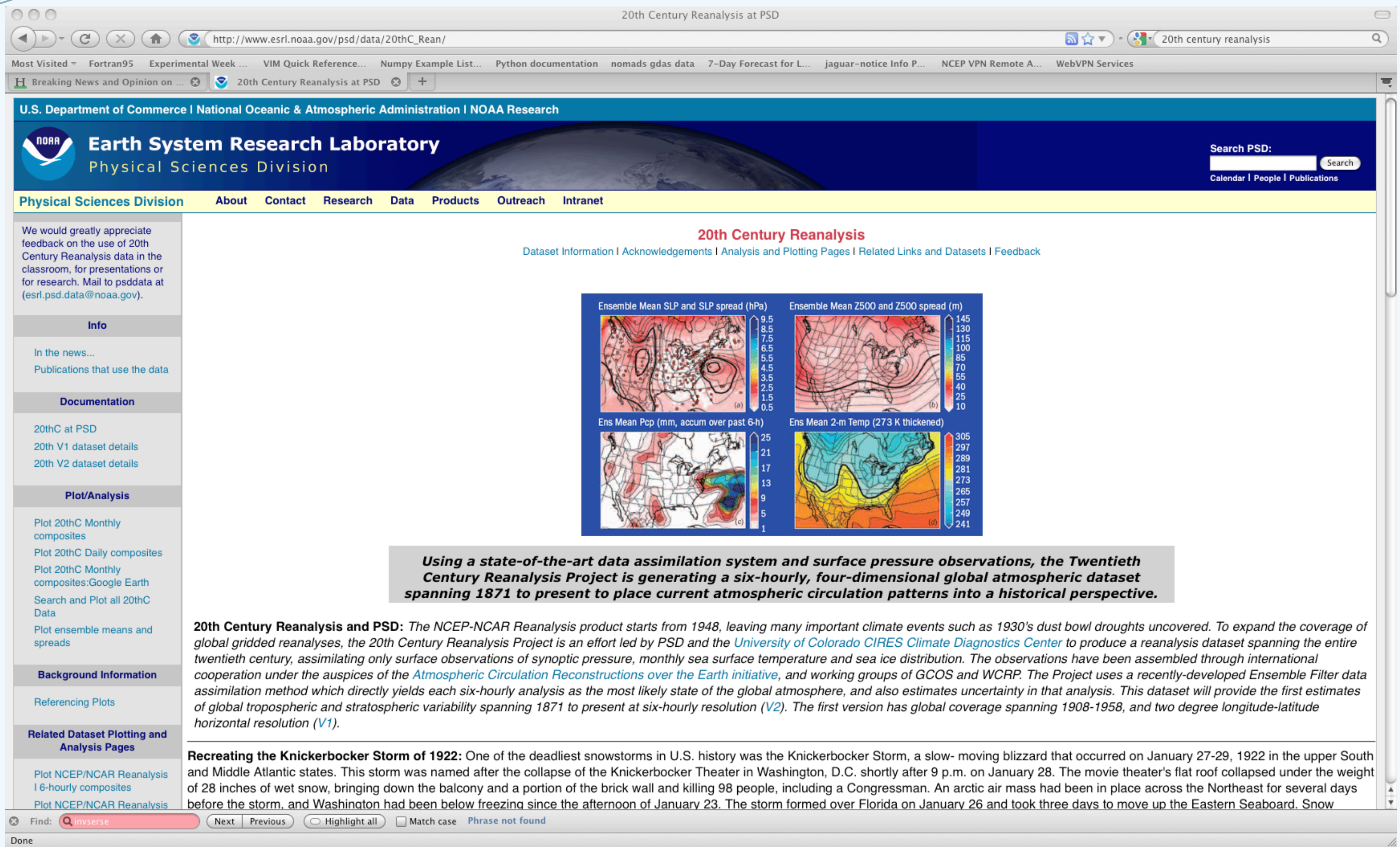
(20th Century Reanalysis Project, led by Gil Compo)





EnKF accurately captures changing uncertainty as observing network changes.

www.esrl.noaa.gov/psd/data/20thC_Rean



20th Century Reanalysis at PSD

http://www.esrl.noaa.gov/psd/data/20thC_Rean/

U.S. Department of Commerce I National Oceanic & Atmospheric Administration I NOAA Research

Earth System Research Laboratory
Physical Sciences Division

Physical Sciences Division | About | Contact | Research | Data | Products | Outreach | Intranet

We would greatly appreciate feedback on the use of 20th Century Reanalysis data in the classroom, for presentations or for research. Mail to psddata at (esrl.psd.data@noaa.gov).

Info
- In the news...
- Publications that use the data

Documentation
- 20thC at PSD
- 20th V1 dataset details
- 20th V2 dataset details

Plot/Analysis
- Plot 20thC Monthly composites
- Plot 20thC Daily composites
- Plot 20thC Monthly composites:Google Earth
- Search and Plot all 20thC Data
- Plot ensemble means and spreads

Background Information
- Referencing Plots

Related Dataset Plotting and Analysis Pages
- Plot NCEP/NCAR Reanalysis I 6-hourly composites
- Plot NCEP/NCAR Reanalysis

20th Century Reanalysis

Dataset Information I Acknowledgements I Analysis and Plotting Pages I Related Links and Datasets I Feedback

Using a state-of-the-art data assimilation system and surface pressure observations, the Twentieth Century Reanalysis Project is generating a six-hourly, four-dimensional global atmospheric dataset spanning 1871 to present to place current atmospheric circulation patterns into a historical perspective.

20th Century Reanalysis and PSD: *The NCEP-NCAR Reanalysis product starts from 1948, leaving many important climate events such as 1930's dust bowl droughts uncovered. To expand the coverage of global gridded reanalyses, the 20th Century Reanalysis Project is an effort led by PSD and the University of Colorado CIRES Climate Diagnostics Center to produce a reanalysis dataset spanning the entire twentieth century, assimilating only surface observations of synoptic pressure, monthly sea surface temperature and sea ice distribution. The observations have been assembled through international cooperation under the auspices of the Atmospheric Circulation Reconstructions over the Earth initiative, and working groups of GCOS and WCRP. The Project uses a recently-developed Ensemble Filter data assimilation method which directly yields each six-hourly analysis as the most likely state of the global atmosphere, and also estimates uncertainty in that analysis. This dataset will provide the first estimates of global tropospheric and stratospheric variability spanning 1871 to present at six-hourly resolution (V2). The first version has global coverage spanning 1908-1958, and two degree longitude-latitude horizontal resolution (V1).*

Recreating the Knickerbocker Storm of 1922: One of the deadliest snowstorms in U.S. history was the Knickerbocker Storm, a slow- moving blizzard that occurred on January 27-29, 1922 in the upper South and Middle Atlantic states. This storm was named after the collapse of the Knickerbocker Theater in Washington, D.C. shortly after 9 p.m. on January 28. The movie theater's flat roof collapsed under the weight of 28 inches of wet snow, bringing down the balcony and a portion of the brick wall and killing 98 people, including a Congressman. An arctic air mass had been in place across the Northeast for several days before the storm, and Washington had been below freezing since the afternoon of January 23. The storm formed over Florida on January 26 and took three days to move up the Eastern Seaboard. Snow

Future Directions

- Operational & prototype product development:
  - Hybrid 4D-Var/EnKF system with NASA/GMAO and NCEP/EMC.
  - EnKF-based hurricane ensembles in "HFIP" program.
  - Higher-resolution 20th-century climate reanalyses

- Research & development issues
  - Unifying global and regional model code
  - Coupled ocean-atmosphere-land-chemistry EnKF reanalyses.
  - Model uncertainty and sampling error in EnKF

Acknowledgments

- We have benefited from collaborations with many university and lab colleagues, including:
  - Eugenia Kalnay, Takemasa Miyoshi, and Brian Hunt (U. MD);
  - Istvan Szunyogh and Liz Satterfield (Texas A&M)
  - Eric Kostelich (University of Arizona)
  - Fuqing Zhang (Penn State)
  - Ryan Torn (U. Albany)
  - Greg Hakim (U. Washington)
  - Xuguang Wang and Ming Xue (U. Oklahoma)
  - Chris Snyder and Jeff Anderson (NCAR)
  - Christian Keppenne (NASA)
  - Peter Houtekamer, Hersh Mitchell, and Mark Buehner (Canadian Meteorological Centre)
  - Craig Bishop, Carolyn Reynolds, and Jim Hansen (NRL)
  - Tomi Vukicevic, Altug Aksoy, and Sim Aberson (AOML)
  - ESRL colleagues Xue Wei, Phil Pegion, Mike Fiorino , Zoltan Toth and Nobuki Matsui

Selected References

- Hamill, T. M., and C. Snyder, 2000A hybrid ensemble Kalman filter / 3D-variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.
- Hamill, T. M., Whitaker, J. S., and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776-2790.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913-1924.
- Whitaker, J. S., G. P. Compo, X. Wei, and T. M. Hamill, 2003: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.* , **132**, 1190-1200.
- Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation Chapter 6 of *Predictability of Weather and Climate*, Cambridge Press, 124-156.
- Compo,G.P., J.S. Whitaker, and P.D. Sardeshmukh, 2006: Feasibility of a 100 year reanalysis using only surface pressure data. *Bull. Amer. Met. Soc.*, 87, 175-190.
- Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II: real observation experiments. *Mon. Wea. Rev.*, **136**, 5132-5147.
- Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. Mon. Wea. Rev., **136**, 463-482.
- Whitaker, J.S., G.P. Compo, and J.N. Thépaut, 2009: A Comparison of Variational and Ensemble-Based Data Assimilation Systems for Reanalysis of Sparse Observations. *Mon. Wea. Rev.*, **137**, 1991–1999.
- Wang, X., T. M. Hamill, J. S. Whitaker, C. H. Bishop, 2009: A comparison of the hybrid and EnSRF analysis schemes in the presence of model error due to unresolved scales. *Mon. Wea. Rev.*, **137**, 3219-3232.

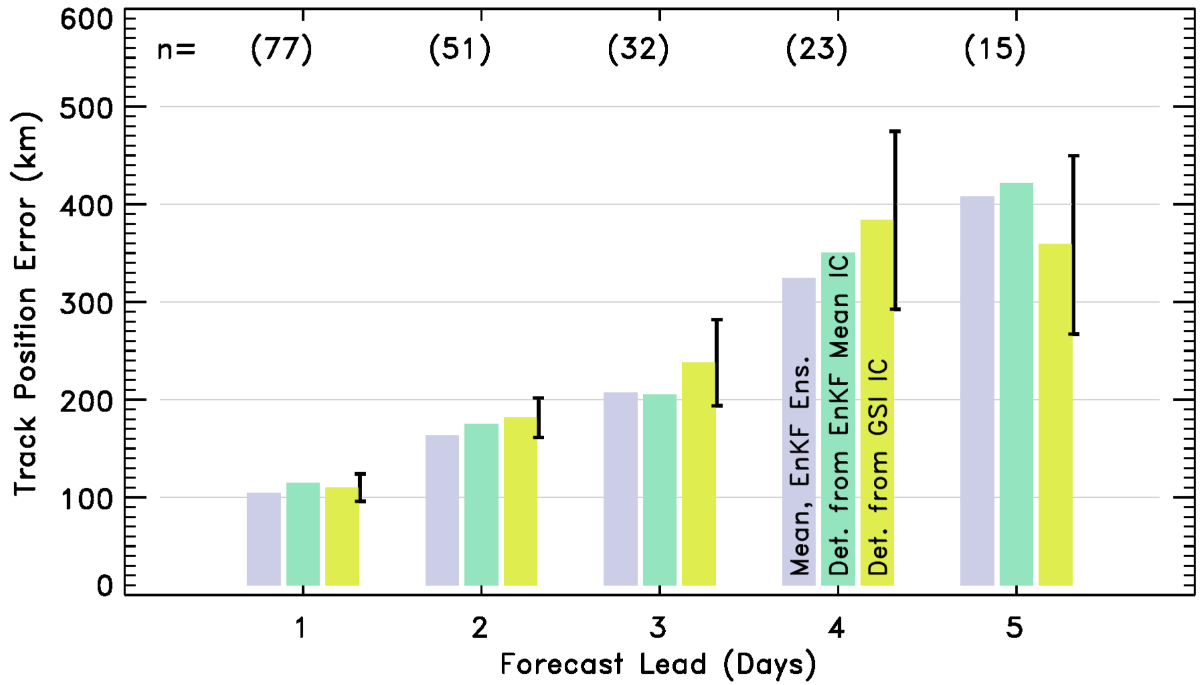
T382 deterministic track forecasts from

(1) mean of EnKF ensemble

(2) deterministic from EnKF mean initial condition, and

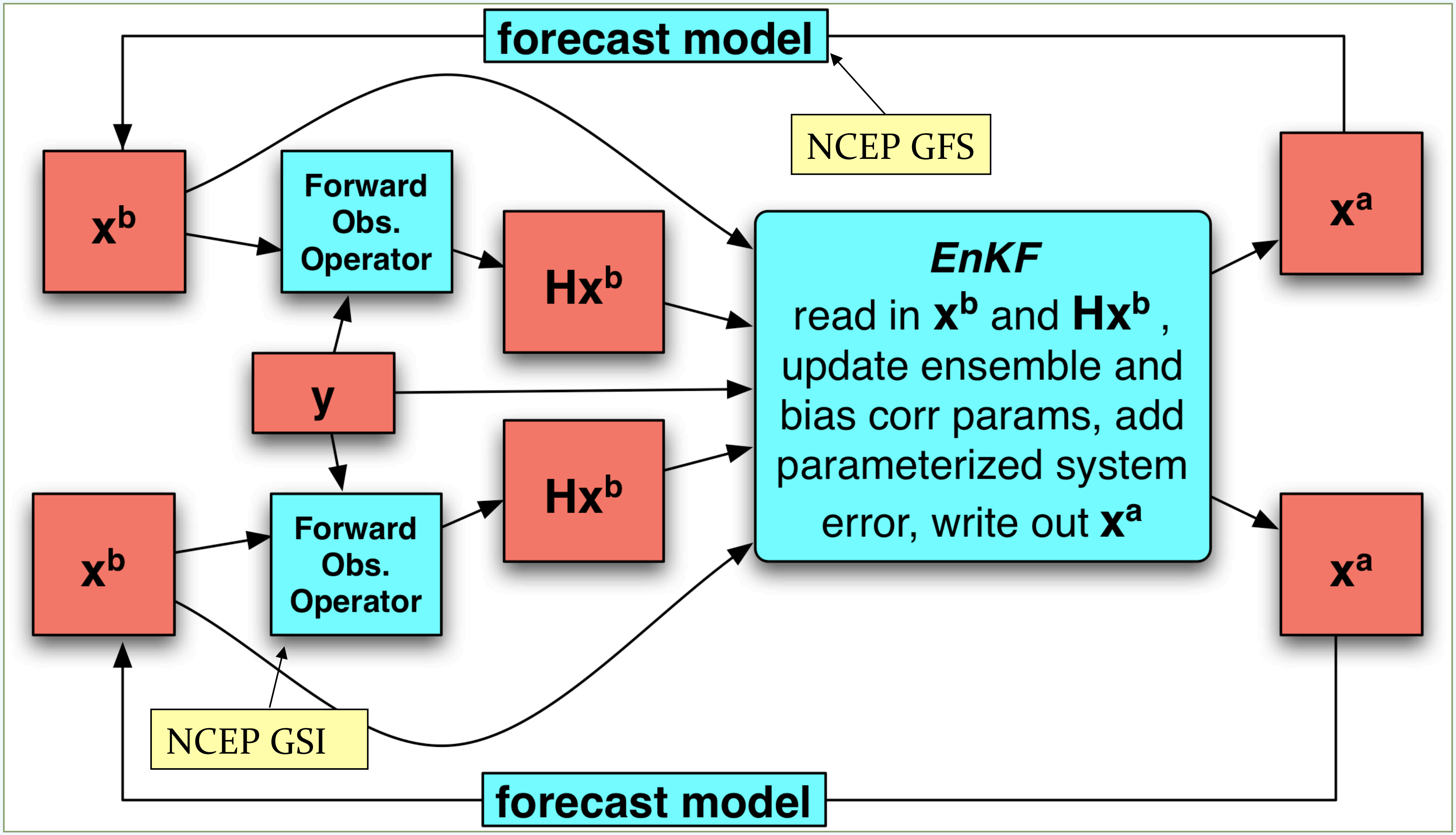
(3) deterministic from "parallel" T382 GFS/GSI

This "parallel" version of GSI now uses TCVitals min pressure pseudo-obs



General slight improvement of EnKF except at longest lead (small sample).

How the NOAA EnKF Works



The Deterministic Serial EnKF – A Recipe

Given a single ob y^o with expected error variance R, an ensemble of model forecasts $\mathbf{x}^b$ (model priors), and an ensemble of predicted observations $\mathbf{y}^b = \mathbf{H}\mathbf{x}^b$ (observation priors):

Step 1: Update observation priors.

(1a) $\overline{\mathbf{y}}_a = (1 - K)\overline{\mathbf{y}}_b + K y^o$ *update for ob prior means*

(1b) $\mathbf{y}'_a = \sqrt{(1 - K)}\,\mathbf{y}'_b$ *rescaling of ob prior perturbations*

where the scalar K = $\mathrm{var}(\mathbf{y}^b)/(\mathrm{var}(\mathbf{y}^b)+ R)$, overbar denotes means, prime denotes perturbations, superscript *b* denotes prior, *a* denotes analysis.

Linear interpolation between observation and observation prior mean *with weight K (0<=K<=1), rescaling of observation prior ensemble so posterior variance is consistent with Kalman filter, i.e.* $\mathrm{var}(\mathbf{y}^a)=(1-K)\,\mathrm{var}(\mathbf{y}^b)= \mathrm{var}(\mathbf{y}^b)R/(\mathrm{var}(\mathbf{y}^b)+R)$.

when $\mathrm{var}(\mathbf{y}^b) << R$, all weight given to prior.

when $\mathrm{var}(\mathbf{y}^b) >> R$, all weight given to observation.

The Serial EnKF – A Recipe (2)

Step 2: Update model priors.

Let $\Delta\mathbf{x} = \mathbf{x}^a - \mathbf{x}^b$ be analysis increment for model priors, $\Delta\mathbf{y} = \mathbf{y}^a - \mathbf{y}^b$ is analysis increment for observation priors.

(2) $\Delta\mathbf{x} = \mathbf{G}\Delta\mathbf{y}$ *computation of increments to model prior*

where $\mathbf{G} = \mathrm{cov}(\mathbf{x}^b, \mathbf{y}^{bT})/\mathrm{var}(\mathbf{y}^b)$

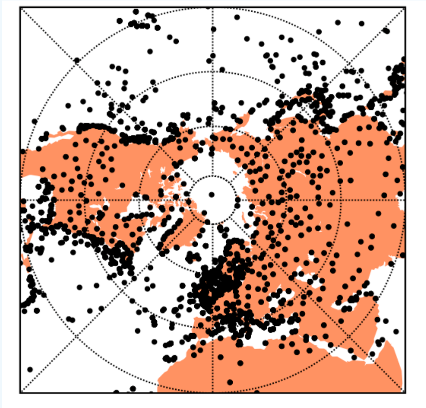
Linear regression of model priors on observation priors.

Only changes model priors when $\mathbf{x}^b$ and $\mathbf{y}^b$ are correlated within the ensemble.

If there is more than one ob to be assimilated, the observation priors for other obs ($\mathbf{Y}^b$) should be also be updated using (2) with $\Delta\mathbf{x}$ replaced by $\Delta\mathbf{Y}$. Next iteration, replace $\mathbf{y}^b$ with next column of $\mathbf{Y}^b$, removing that column from $\mathbf{Y}^b$. After each iteration the model priors and observation priors are set to the latest analysis values ($\mathbf{x}^a$ replaces $\mathbf{x}^b$, $\mathbf{Y}^a$ replaces $\mathbf{Y}^b$). Continue iterating until $\mathbf{Y}^b$ is empty.

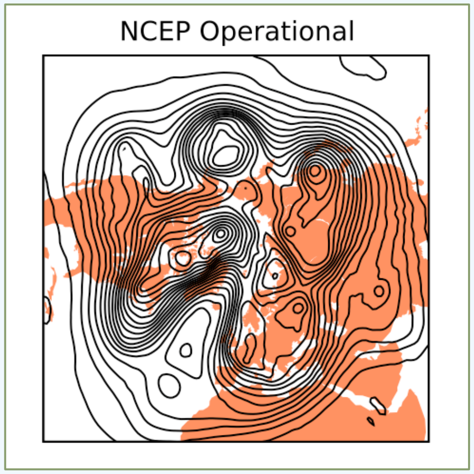
Example : 500-hPa Height Analyses Assimilating Only p_s obs

Whitaker et al 2009: June MWR

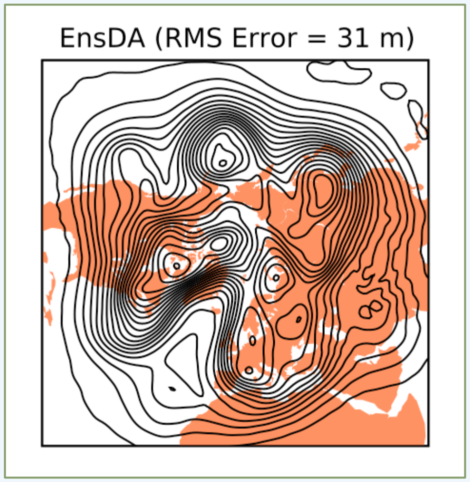


Black dots show 300+ surface pressure observation locations (similar to 1930's network)

Full NCEP operational analysis (3D-Var)
1000000+ obs



EnKF
only 300+surface pressure obs



ECMWF 3D-Var
only 300+surface pressure obs